MEASURING EDUCATIONAL PROCESSES THROUGH EDUCATIONAL RESULTS[1]

LEONARD P. AYRES
Director of the Division of Education, The Russell Sage Foundation

Fifteen years ago the school superintendents of America, assembled in convention in Indianapolis, discussed the problems then foremost in educational thought and action. At that meeting a distinguished educator—the pioneer and pathfinder among the scientific students of education in America—presented the results of his investigations of spelling in the school systems of nineteen cities. These results showed that, taken all in all, the children who had spent forty minutes a day for eight years in studying spelling did not spell any better than the children in the schools of other cities where they devoted only ten minutes per day to the study.

The presentation of these data threw that assemblage into consternation, dismay, and indignant protest. But the resulting storm of vigorously voiced opposition was directed, not against the methods and results of the investigation, but against the investigator who had pretended to measure the results of teaching spelling by testing the ability of the children to spell.

In terms of scathing denunciation the educators there present, and the pedagogical experts, who reported the deliberations of the meeting in the educational press, characterized as silly, dangerous, and from every viewpoint reprehensible the attempt to test the efficiency of the teacher by finding out what the pupils could do. With striking unanimity they voiced the conviction that any attempt to evaluate the teaching of spelling in terms of the ability of the pupils to spell was essentially impossible and based on a profound misconception of the function of education.

Last week, in the city of St. Louis, that same association of school superintendents, again assembled in convention, devoted

[1] An address delivered before the Harvard Teachers' Association, Cambridge, Massachusetts, March 9, 1912.

forty-eight addresses and discussions to tests and measurements of educational efficiency. The basal proposition underlying this entire mass of discussion was that the effectiveness of the school, the methods, and the teachers must be measured in terms of the results secured.

This change represents no passing fad or temporary whim. It is permanent, significant, and fundamental. It means that a transformation has taken place in what we think as well as in what we do in education. It means that education is emerging from among the vocations and taking its place among the professions.

This profound change in our educational practice has not come through the slow processes of philosophy, nor because we were awakened by the stirring words of voice or pen of any educational prophet. Few school men can claim great credit for having hastened its advent. It was forced upon us, first, by the natural results of compulsory education, and still more definitely and directly by the exactions of the scientific age in which we live.

THE SCIENTIFIC METHOD IN EDUCATION AND INDUSTRY

This new attitude of educators toward education means that we have ceased exalting the machinery and have commenced to examine the product. We have awakened to a startled realization that in education, as in other forms of organized activity, applied science may avail in improving even those processes that have rested secure in the sanction of generations of acceptance.

The transformation now taking place in education means that it is our privilege to be part of a movement that is working changes comparable to those that are now remaking almost every form of industrial activity. The trade of brick-laying, practiced by millions of intelligent artisans, has remained almost unaltered since the days of primitive man. But scientific management steps in and asks, Why lower a hundred pounds of human flesh to pick up each two-pound brick? Why toss the brick four times to find its best face? Why tap it three times to get its proper level? Why stand in a position that requires a half-dozen movements when one will suffice? And science makes answer: Build a platform for

the bricks adjusted to the height of the work; lay the bricks on the platform with the best face out; mix the mortar so that one tap will suffice; and take such a position that five movements accomplish the same results that formerly required eighteen. The result is that each workman lays as many bricks in one hour as he formerly laid in three.

The ideals and processes of the application of the scientific method to education are in salient respects similar to those that are reshaping the processes of industry. In education, as in industry, the scientific idea is at base analytic scrutiny, exact measuring, careful recording, and judgment on the basis of observed fact. Swiftly, silently, and almost without warning, the scientific methods have invaded the educational camp and have begun to demolish the hosts of theory, legend, superstition, and tradition.

The time has already passed for us to query whether or not we shall indorse and adopt the new scientific attitude of exact measurement and judgment by results. The new method is upon us and the question at issue is no longer, "Shall we adopt it?" but rather, "How shall we utilize it?"

REFORM IN EDUCATIONAL ADMINISTRATION

Three years ago twenty-nine cities in America had systems of individual record cards for keeping the school histories of their children. Today 216 cities have adopted a uniform system for this purpose. Those cities intend to judge processes by results.

One year ago the number of city school systems having uniform records of accounting whereby the school facts of one locality could be compared with those of another was about fifteen. Today the number of such cities is 418. Their aim is a mutual comparison of results.

Seven years ago Superintendent Maxwell, of New York City, published data in his annual report showing that 39 per cent of the school children of that city were above the normal ages for their grades. Judged by the age standards, they were educational misfits. At that time these data were almost unique and attracted widespread attention because of their unusual character. Today such methods of checking up the results of our school work are

commonplace and a few months ago the Federal Bureau of Education published similar data showing conditions in 318 cities.

These nation-wide changes are not products of mere chance. They have come because the public and the educators have begun to demand real information about their public schools. Less than five years ago it occurred to a few people in America seriously to ask the question, "What proportion of the children who enter our common schools remain to complete the course?" This was a plain business proposition. The question at issue was the relation of the finished product to the raw material. The children who enter our public schools in the first grade are the raw material; those who complete the course and graduate are the finished product. It was an elementary question in business administration that these students were asking when they inquired what proportion of the children complete the common-school course.

In order to answer this question we must have two figures: First, the number of children who graduate. That can easily be ascertained in any school system. Second, the number of children who begin school each year. That cannot be obtained so easily. Incredible as it may seem, up to five years ago school men had never thought it worth while to record that datum. A patient search showed that the cities in America recording the number of children entering school each year could be counted on the thumbs of two hands. Today the number of cities keeping such records runs into the scores. At that time, all that the school superintendents knew about the matter was that the beginners were numerous; that progress was not uncommon; and that there were some graduates each year. Now they know that in the country as a whole not one-half of the children who enter the public schools remain to graduate, and they are busily at work remaking their school systems to remedy that condition. The startling revelation that our vaunted system of free education was failing to give even complete elementary schooling to a majority of the children evoked imperious demands for more real facts. Here were statements of educational conditions within the comprehension of all and painfully obvious in their significance. They left no room for question as to the necessity for checking up results in education.

The school children are the invested capital of the community. What should we say of a bank that kept its accounts in the same way that the school has kept account of the invested capital of society? What would you say if your banker should confess that the only facts revealed by his books were the total number of accounts handled during the year and the average monthly assets? What would you say if he should confess that he did not know and could not find out anything about the number or amount of new accounts received, old ones withdrawn, or the results of his investments?

Nor was this situation confined to elementary schools. Conditions in our higher schools were even more shocking. The proportion of the pupils who completed the course as compared with those who entered was startlingly small. We can hardly imagine an analogous situation in any commercial industry. What, for example, should we say of a four-act play in a theater where a thousand people were present at the beginning of the first act, five hundred got up and left before the beginning of the second act, two hundred and fifty of these refused to sit through the third, and only one hundred and twenty-five remained to see the final descent of the curtain? And yet these figures express the situation in many of our larger cities with respect to the falling out of pupils in the four years of our high-school course.

AUTHORITY VERSUS EVIDENCE

The new method which judges processes in terms of results has been by no means confined to the development of record forms and the perfecting of new devices in the statistics of school administration.

About three years ago a graduate student in one of the universities of Massachusetts tried to investigate the old problem, "What is the best age at which to send a child to school?" In his search for information he put the question to the head of every college department of pedagogy in the country. He received definite and positive replies from almost all to the effect that the best entering age is a comparatively late one. He then followed his first inquiry by a second in which each pedagogical expert was

asked on what he based his assertion. In every case save one the answer was that the writer was positive of the correctness of his views, but had no evidence with which to substantiate them. The exception was a man who said that he knew, because his own son had entered school late and had made good progress. This happened only three years ago and the answers were speculative and indefinite because quantitative evidence bearing on the problem did not exist. And yet so rapid is the progress that has since been made that there is published in the current number of *Education* a study of that problem based on the school histories of more than twenty-five thousand children.

From a Michigan city there comes a striking illustration of the degree to which educators have enjoyed that freedom which comes through being entirely unhampered by facts. About five years ago a movement gained headway in that city for the establishment of kindergartens. The advocates of the proposed innovation gave as their most weighty argument the claim that children who pass through the kindergarten subsequently complete the elementary grades in less time than do those who have not enjoyed the advantages of such training. The faction opposing the establishment of the kindergartens denied the validity of this argument. To settle the question the school authorities wrote to school superintendents all over the country asking whether children who had gone through the kindergartens subsequently completed the work of the grades more quickly than did those who had not received such training. Replies were received from the superintendents of 72 cities. Of these, 49 answered that they thought that children having kindergarten training subsequently made more rapid progress than the others, but that they did not know. The other 23 replied that they held the opposite opinion, but that they did not know.

That result was typical of the supremacy of speculation over evidence in education. In this problem, as in other problems, opinions have varied. There has been a consensus of belief but there has been an almost absolute absence of definite knowledge. Kindergartens have been increasingly numerous in America since Elizabeth Peabody established the first one in Boston in 1868.

They now exist by the thousands, and on them we have spent each year hundreds of thousands of dollars. During the entire period a favorite argument in their support has been the one relied on to secure their establishment in the Michigan city, and yet until the past five years no one has been able to state in definite terms anything about the real effect of kindergarten training. This situation no longer exists, for within the past three years extensive investigations have been conducted, comparing the school histories of many thousands of children who have had kindergarten training with the school histories of the children in the same systems who have not had kindergarten training.

EDUCATIONAL SURVEYS

The new scientific method has not been confined to the investigation of isolated problems. In city after city across the country its aid is being invoked to evaluate educational results through the medium of the school survey. Unheard of only a few years ago, these city-wide educational inquiries have been made or are in progress in such cities as Boston, Baltimore, Boise, Montclair, Orange, and New York. Already they are being planned in other localities, and one embracing the system of the entire state of Wisconsin is now under way.

CONSERVATISM VERSUS PROGRESSIVISM

The progress of this educational revolution has been stoutly contested and each forward step has been greeted by an anvil chorus of opposition in which the notes ranged from the grudging admissions of the skeptic to the fiery denunciations of the educational reactionary. Always retiring and always fighting, these forces of opposition have abandoned as untenable their contention of fifteen years ago, that any and all attempts at measurement in education are silly and dangerous. Having given up this position, they next took refuge in the firm declaration that while material matters in education may be quantitatively investigated, the immaterial problems of the teaching process can never be submitted to such treatment. They admitted that it would probably do no harm to discover the more important facts with

respect to financial expenditures and the progress of pupils, but firmly declared that no phase of intellectual phenomena would yield to statistical analysis.

THE MEASUREMENT OF EDUCATIONAL PRODUCTS

No sooner was this doctrine fully formulated than there appeared a set of scientific students of education presenting measuring scales with which to gauge the performance of the children in their classroom work. Thorndike with his measuring scales for handwriting, Stone and Courtis with their standardized tests in arithmetic, and Hillegas with his method for measuring the quality of English composition again forced the champions of tradition to retire and find a new point of defense.

CHARACTER AND EFFICIENCY

The final citadel in which the old guard is now making its last stand consists of the objection that the most important elements of true teaching can never be measured.

They claim, and they are right in claiming, that we can never determine by mathematics the degree to which the strong man and the noble woman can influence for good the characters of their pupils. But what they overlook is the fundamental truth that in education, as in other pursuits of life, character and efficiency go hand in hand. As school executives make practical application of the newer scientific tests, no fact stands out with more impressive distinctness than that the teachers whose classes make the best records are the teachers who are the most truly successful in the shaping of character.

INDIVIDUAL DEVELOPMENT NOT UNIVERSAL UNIFORMITY

There remains one other objection, less frequently advanced, but sometimes voiced, namely, that the advocates of the scientific method aim to reduce all work in education to the dead level of uniform precision. This charge is born of a complete misunderstanding of the ends, aims, and processes of the new method. Its aim is not uniformity but individual development. The measured beat of the concert recitation is not music to the ears of the scientific students of education. The sight of a rigid row of reciting

children with toe tips nicely adjusted to a line painted on the classroom floor does not cause their souls to leap in admiration. Their ideal of school discipline does not consist of having a roomful of growing children accomplish the amazing feat of sitting through an entire period without moving a muscle or winking an eye. Their ideal of educational administration does not contemplate a uniform country-wide daily program by which each recitation period in every city and hamlet shall be fixed by a master clock located at the seat of the national government in Washington.

THE SCIENTIFIC METHOD MEANS THE MEASUREMENT OF RESULTS

The object of the new method is the substitution of evidence for opinion and knowledge for speculation. Its champions are working to develop measurements in education because they realize that only by this method can education become an art and a science and its practice be changed from a vocation to a profession. They scan the history of science and remember that through the development of measurements astronomy grew out of astrology, chemistry emerged from alchemy, and physics developed from mystery.

They read the history of education and realize that the astonishing progress of the past decade has come from shifting the position of inquiry from asking "What results can or might we get?" to "What results are we getting?" This makes the pupil and not the teacher the center of interest. It calls a halt on the futile quest for standards of attainment on which we have never come to an agreement, and aims instead to discover units of measurement. Simple as it sounds, this change from asking "What results should we get?" to asking "What results are we getting?" is the keynote of the whole scientific method in education. To answer the question in its new form means the development of units of measurement, and when these are secured the standards of attainment will work themselves out automatically.

THE FUTURE

The certainty about the scientific method in education is that it is with us. That it will develop enormously in the immediate

future is entirely sure. What its effects will be we can as yet only surmise. The dangers involved are as real and imminent as the advantages are self-evident. These dangers will arise from the mass of superficial and erroneous results that will certainly be presented to the educational world in the guise of scientific contributions to applied pedagogy. What is to be our attitude toward each new contribution?

My own answer is that we must welcome them all, but challenge them all, and attempt to verify them all. Every figure, every process, and every conclusion, whether presented by the educational expert or by the novice, must be submitted to the most rigid scrutiny and searching analysis before being accepted as worthy of inclusion in the new pedagogy.

In proportion as we are thus enabled to retain the genuine and reject the spurious, education will move forward among the other sciences. Its new methods will substitute knowledge for speculation and evidence for opinion. Its marshaled facts, expressed in definite terms, will demolish the hosts of legend, superstition, tradition, and theory.

Under the new régime the studies to be included in the curriculum and the methods by which they are taught must have a more valid reason for being than the fact that our forefathers had them in their schools.

"How much?" and "How many?" and "With what result?" are going to displace guess-work, imagination, and oratory as criteria for shaping educational policies. The old method has been education within the sheltering walls of the cloister in which an occasional peep-hole has been cut, to satisfy the parent and silence the taxpayer. The new method proposes education in the open and under the clear and penetrating rays of the search-light.